REMARKS

Claims 1-19 are pending in this application, of which claims 1, 17, and 18 are independent. Favorable reconsideration of the final Office Action mailed October 6, 2009, is respectfully requested in view of the foregoing amendments and the following remarks.

Interview Summary

On December 7, 2009, Examiner Greg Borsetti and the applicant's representative, Mandy Jubang of Occhiuti Rohlicek & Tsao LLP, conducted a telephone interview. The claim language set forth in claim 1 was discussed in view of the Cardillo and Wolf references of record. The examiner and the applicant's representative agreed that additional language related to receipt of input from a user identifying portions of a first set of audio signals being of interest to the user and use of such input in generating subword unit representations would clarify the claim and more clearly distinguish the Cardillo and Wolf references of record.

Claim Objections

Claim 13 has been amended to include a space between "claim" and "1".

35 U.S.C. § 112, Second Paragraph, Rejections

Claims 17 and 18 have been amended to provide the requisite antecedent basis for the term "unknown speech." Withdrawal of the 35 U.S.C. § 112, second paragraph rejections is respectfully requested.

35 U.S.C. § 103 Rejections

Claims 1-4, 8, 9, and 12-19 are rejected as being unpatentable over Cardillo et al., "Phonetic Searching vs. LVCSR: How to Find What You Really Want in Audio Archives", in view of Wolf et al., U.S. Patent Application Publication No. 2003-0204492. Claims 5-7, 10, and 11 are rejected as being unpatentable over Cardillo et al., in view of Wolf et al., and further in view of Ferrieux et al., "Phoneme-Level Indexing for Fast and Vocabulary-Independent Voice/Voice Retrieval".

Amended claim 1 requires, in part, "receiving input from a user identifying at least two portions of a first set of audio signals as being of interest to the user; processing, by a query recognizer of a word spotting system, each identified portion of the first set of audio signals to generate a corresponding subword unit representation of the identified portion; [and] forming, by the query recognizer of the word spotting system, a representation of a spoken event of interest, wherein the forming includes combining the subword unit representations of the respective identified portions of the first set of audio signals."

Cardillo teaches searching digital audio at a word or phrase level. Specifically, the searching phase described on page 12 of Cardillo starts with processing a text-based single- or multi-word query to form a phonetic representation of the text-based query. (Table 1 of Cardillo shows single-word query terms (e.g., add, age), two-word query terms (e.g., nothing but, different was), three-word query terms (e.g., theater missile defense, the nasdaq index), and four-word query terms (e.g., president elect bush said, balance clarity and depth) and their respective number of phonemes.) Next, a phonetic search track representative of the digital audio is phonetically searched using the phonetic representation of the text-based single- or multi-word query.

Cardillo does not disclose "receiving input from a user identifying at least two portions of a first set of audio signals as being of interest to the user." At most, Cardillo teaches receiving an indication that a text-based single- or multi-word term is of interest to a user. Further, Cardillo does not disclose "processing... each identified portion of the first set of audio signals to generate a corresponding subword unit representation of the identified portion; [and] forming... a representation of a spoken event of interest, wherein the forming includes combining the subword unit representations of the respective identified portions of the first set of audio signals." At most, Cardillo teaches forming a representation of a text-based query by probing a phonetic dictionary and/or consulting a spelling-to-sound database. See, e.g., page 12 of Cardillo: "A phonetic dictionary is probed for each word within the query term to accommodate unusual terms (whose pronunciations must be handled specially for the given natural language) as well as very

common words (for whom performance optimization is worthwhile). Any word not found in the dictionary is then processed by consulting a spelling-to-sound data base to extract likely phonetic representations given the word's orthography."

Wolf discloses retrieving documents from a multimedia database using spoken queries. In paragraph 0034, Wolf states: "A spoken query 105 to search 180 the database 140 is processed by the search engine 190 as follows. The spoken query is provided to the speech recognition engine 150. However, instead of converting the spoken query directly to text, as in the prior art, the system according to the invention generates a lattice 106. In the lattice 106, the nodes represent the spoken words, and the directed edges connecting the words represent orders in which the words could have been spoken. Certainty information is retained with the nodes and edges. Generally, the certainty information includes statistical likelihoods or probabilities. Thus, the lattice retains the certainty due to ambiguities in the spoken query."

Even if, for the sake of argument only, Wolf's "spoken query" is read as corresponding to the recited "spoken event of interest," and even if the audio signals corresponding to the spoken query that is provided to the speech recognition engine is read as corresponding to the recited "first set of audio signals," no portion of Wolf provides any hint or disclosure of "receiving input from a user identifying at least two portions of a first set of audio signals as being of interest to the user," much less "processing... each identified portion of the first set of audio signals to generate a corresponding subword unit representation of the identified portion; [and] forming... a representation of a spoken event of interest, wherein the forming includes combining the subword unit representations of the respective identified portions of the first set of audio signals," as required in amended claim 1.

For at least these reasons, the Applicant respectfully submits that Cardillo, whether taken alone or in any proper combination with Wolf, does not describe or suggest all of the features of amended claim 1.

The dependent claims 2-16 are patentable for at least similar reasons as the claims on which they depend are patentable.
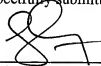
Applicant(s) : Robert W. Morris   Attorney Docket No.: 30004-004US1
Serial No. :   10/565,570
Filed   :   July 21, 2006
Page   :   10 of 10

The independent claims 17 and 18 are patentable for at least similar reasons given above for claim 1.

Conclusion

It is believed that all of the pending claims have been addressed. However, the absence of a reply to a specific rejection, issue or comment does not signify agreement with or concession of that rejection, issue or comment. In addition, because the arguments made above may not be exhaustive, there may be reasons for patentability of any or all pending claims (or other claims) that have not been expressed. Finally, nothing in this paper should be construed as an intent to concede any issue with regard to any claim, except as specifically stated in this paper, and the amendment of any claim does not necessarily signify concession of unpatentability of the claim prior to its amendment.

A Request for Continued Examination is being filed with this amendment. The Request for Continued Examination fee in the amount of $405 is being paid concurrently herewith on the Electronic Filing System (EFS) by way of Deposit Account authorization. Please apply any other charges or credits to Deposit Account No. 50-4189, referencing Attorney Docket No. 30004-004US1.

Respectfully submitted,

Date: 12/8/09

Mandy Jubang
Reg. No. 45,884

Customer No. 69713
Occhiuti Rohlicek & Tsao LLP
10 Fawcett Street
Cambridge, MA 02138
Telephone: (617) 500-2504
Facsimile: (617) 500-2499